DOCUMENT RESUME

ED 328 622 UD 027 869

AUTHOR	Gamoran, Adam
TITLE	The Variable Effects of Tracking: Inequality and Productivity in American High Schools.
INSTITUTION	National Center on Effective Secondary Schools, Madison, WI.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Oct 90
CONTRACT	OERI-G-008690007-90
NOTE	40p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Ability Grouping; *Academic Achievement; *Catholic Schools; *High Schools; National Surveys; *Public Schools; *Track System (Education)
IDENTIFIERS	*Hierarchical Linear Modeling; High School and Beyond (NCES)

ABSTRACT

The effects of tracking on high school students' academic achievement are not constant across schools. This study examines the relationship of the following organizational dimensions of tracking systems on the variation in between-track inequality and productivity: (1) selectivity; (2) electivity; (3) inclusiveness; and (4) scope. "Inequality" refers to the "achievement gap" between tracks; "productivity" refers to the average achievement of each school. In addition, differences between public and Catholic schools are examined. Statistical data from the High School and Beyond survey were analyzed for 28,804 students in 964 public and Catholic high schools. The following findings are reported: (1) more flexible systems produce lower inequality in mathematics, reading, and vocabulary achievement; (2) moderately inclusive systems also produce lower inequality in mathematics achievement; (3) flexibility and inclusiveness have a positive effect on overall achievement; (4) elective systems produce higher achievement; and (5) Catholic schools have lower inequality and higher productivity than public schools, due in part to the way they use tracking. A list of 59 references and 5 tables of statistical data are appended. (FMW)

DRAFT--COMMENTS WELCOME

THE VARIABLE EFFECTS OF TRACKING: INEQUALITY AND PRODUCTIVITY IN AMERICAN HIGH SCHOOLS

Adam Gamoran
University of Wisconsin, Madison

October 1990

Earlier versions of this paper were presented at the Centre for Educational Sociology in Edinburgh, Scotland, and at the 1988 meetings of the American Sociological Association in Atlanta. The author is grateful for comments received at those forums, and for the helpful suggestions of Anthony Bryk, Robert Hauser, Michael Olneck, and Douglas Willms. The author also appreciates the assistance of Dae-Dong Hahn in preparing the data set, and of the staff of the Social Science Computing Cooperative, particularly Pam Stangl and Mary Rasmussen, in implementing the HLM computer program. Research on this paper was supported by the National Center on Effective Secondary Schools at the Wisconsin Center for Education Research, which is supported by a grant from the Office of Educational Research and Improvement (Grant No. G-008690007-90). The final draft was prepared with the support of a Spencer Fellowship to the National Academy of Education. Any opinions, findings, and conclusions expressed in this paper are those of the author and do not necessarily reflect the views of these agencies or the U.S. Department of Education.

THE VARIABLE EFFECTS OF TRACKING:
INEQUALITY AND PRODUCTIVITY IN AMERICAN HIGH SCHOOLS

ABSTRACT

This paper suggests that the effects of high school tracking on student achievement are not constant across schools. Based on the work of Sorensen (1970), the author describes four structural dimensions of tracking systems: selectivity, electivity, inclusiveness, and scope. Variation in these organizational conditions are predicted to lead to variation in between-track inequality (the achievement gap between tracks) and productivity (average achievement in the school), net of the composition of the student body. In addition, Catholic schools are predicted to have lower inequality and higher productivity than public schools. The hypotheses are tested with data on from the nationally representative High School and Beyond survey. The results indicate that more flexible tracking systems produce lower inequality in math, reading, and vocabulary achievement. In math, moderately inclusive systems also reduce inequality. Overall achievement is positively affected by flexibility and inclusiveness. In vocabulary, schools with more students who believe they chose their own track score higher as well. The hypotheses about Catholic schools are also supported, and the way Catholic schools use tracking accounts for part of their advantage.

THE VARIABLE EFFECTS OF TRACKING: INEQUALITY AND PRODUCTIVITY IN AMERICAN HIGH SCHOOLS

Microsocial processes often vary according to the larger contexts in which they occur. When researchers explicitly look for it, such contextual variation may be detected. For example, the influence of schooling on the number of children a woman bears depends on the presence of family planning programs in specific countries (Mason, Wong, and Entwistle, 1983). Similarly, the relation between students' family backgrounds and their likelihood of finishing high school varies according to the demographic characteristics of their schools (Bryk and Thum, 1989). In both examples, the effects of a micro-level process (e.g., schooling on fertility) depend on the presence of a macro-level condition (a national family-planning program). Despite the added complexity brought on by modeling such context-dependence, appropriate statistical techniques have become available to sociologists (e.g., Mason, Wong, and Entwistle, 1983; Bryk and Raudenbush, 1986). Yet theoretical development, which is also made more complex by considering multi-level contextual variation, has not always kept pace with the methodological advances. Few topics have been explored for the presence of multi-level contextual effects.

For instance, many writers have noted the possibility that the effects of high school tracking on achievement may vary between schools (e.g., Heyns, 1974; Hauser, Sewell, and Alwin, 1976; Rosenbaum, 1984), but no compelling theory has been presented to predict when such effects might be large and when they are small. Do different kinds of tracking systems have different effects on achievement? If so, why? To the extent that track effects do vary between schools, analyses of tracking that assume otherwise would be misleading (e.g., Gamoran, 1987). This paper uses existing knowledge about tracking and its effects to develop hypotheses about between-school differences in the effects of tracking. Building on the work of Sorensen (1970), it argues that track effects differ according to certain structural characteristics of stratification systems, which vary across schools. The paper also assess previous speculation that the effects of tracking are not the same in Catholic schools as they are in public schools (Gamoran and Berends, 1987). These hypotheses are tested by applying methods designed for multi-level contextual analysis to data on tracking and achievement in a national sample of high schools.

Sorensen (1970) outlined several dimensions of organizational differentiation in school systems. In his scheme, high school tracking is an instance of both horizontal differentiation, involving curricular differences within grade levels, and vertical differentiation, because of status distinctions between academic and non-academic programs (on this point, see also Rosenbaum, 1984, and Lucas, 1990). Not all tracking systems are alike, however. They vary along structural dimensions such as selectivity--the extent of homogeneity within tracks; electivity--the extent to which students choose or are assigned to track positions; inclusiveness--the availability of options for subsequent educational opportunity in a given track system; and scope--the breadth and flexibility of track assignment. These structural conditions have been elaborated by Rosenbaum (1976, 1984) and by Oakes (1985), and I will present them in greater detail shortly, along with a discussion of differences in the tracking systems of public and Catholic schools.

Before predicting how structural and sectoral differences influence the effects of tracking on achievement, however, it is necessary to make clear what is known about the effects of tracking, and about the mechanisms through which they occur. The goal of the next section is to produce a short list of statements about tracking's effects and their sources. These findings will serve as the foundation upon which to build hypotheses concerning between-school variation in track effects.

Tracking and Student Achievement

Tracking may affect achievement in two ways. First, it may affect the dispersion of achievement, or educational inequality. Previous work examined tracking's impact on the dispersion of outcomes across population subgroups (e.g., blacks vs. whites; rich vs. poor) (Gamoran and Mare, 1989). Now, the issue is the dispersion of achievement across the tracks themselves. Tracking adds to inequality when placement in the high-status track permits students to gain more than if they had been assigned to the other track. A key question for this paper is whether some forms of tracking induce more inequality than others.

Second, tracking may influence the overall level of achievement in the school, or educational productivity. This possibility would arise if researchers compared tracking to the absence of tracking (see Gamoran and Mare, 1989), but it is also germane to the question of whether variation in the structure of tracking contributes to between-school variation in achievement. Is one type of tracking system more productive than another? This question must be paired with the study of inequality, because if certain forms of tracking create smaller achievement gaps between tracks, it is essential to know whether the reduced inequality occurs in the context of higher, lower, or the same overall achievement.

Does Tracking Affect Inequality?

For many years, students, teachers, and field researchers have reported that more learning occurs in the higher tracks (Hargreaves, 1967; Rosenbaum, 1976; Metz, 1978; Ball, 1981; Burgess, 1984; Oakes, 1985). Most survey studies corroborate this impression after controlling for students' initial characteristics, including family background, race, gender, and prior achievement (for a review, see Gamoran and Berends, 1987). Although the finding has not been universal (Jencks and Brown, 1975, and Alexander and Cook, 1982, raised doubts), the presence of net track differences in well-specified, carefully controlled analyses using national survey data for Britain, Israel, and the U.S. seems persuasive (Kerckhoff, 1986; Shavit and Featherman, 1988; Natriello, Pallas, and Alexander, 1989). Gamoran and Mare (1989) showed that the finding holds even after taking into account the effects of unmeasured selection variables.

Previous writers have disagreed about whether tracking's differentiating effects vary across schools. Rosenbaum (1984) has been most vehement in arguing that such variation does occur. He attributed differences among survey findings in part to differing track structures, as revealed by several case studies. In a nation-wide study, Oakes (1985) described considerable variation in the characteristics of tracking systems in twenty-five junior and senior high schools, but she did not examine whether these differences affected the impact of tracking on achievement.

Heyns (1974) reported statistically significant interactions between track positions and dummy variables for the high schools in her data set. However, because the interactions were relatively small, and because they were not related to school size or location, she estimated an additive model of track effects. Using the same approach, Hauser, Sewell, and Alwin (1976) found no significant between-school differences in track effects for data from Milwaukee County, Wisconsin. Consequently they, too, estimated an additive model. Both Heyns' data, limited to urban schools outside the south, and Hauser, Sewell, and Alwin's Wisconsin data, might contain less variability than a nationally representative random sample. Because of disagreement over the existence and magnitude of between-school differences in track effects, the present paper will test for homogeneity of effects across schools, before exploring the sources of any differences that appear.

Does Tracking Affect Productivity?

The effects of tracking on educational productivity are less well understood, because although studies of curriculum tracking have compared the different tracks to one another, they have not compared tracked to untracked schools. As Gamoran and Mare (1989) pointed out, tracking's implications for productivity depend on what the alternative is: compared to current practices, a system in which all students encountered a program like the college track would probably produce higher achievement, but if all students were assigned to programs like the noncollege tracks, achievement would likely be lower. In a study of ability grouping in British secondary schools, Kerckhoff (1986) found that compared to students in mixed-ability classes, the gains of high-ability students were offset by losses of low-ability students, so that grouping had little overall effect on productivity.

It has been well established that schools differ from one another in their average achievement levels. That is, even after adjusting for differences in the compositions of their student bodies, some schools appear more productive than others. Although between-school variation is typically less than 20% of the total variation in student achievement, the amount of variation is statistically significant (Bryk and Driscoll, 1988; Bryk and Raudenbush, 1988; Lee and Bryk, 1989). We do not know whether this variation is affected by the structure of tracking in the school. Gamoran (1987) reported higher vocabulary achievement in schools having larger college tracks, but the finding did not hold for math, science, reading, writing, or civics achievement. To date, researchers have not presented a conceptual account for the impact of the structure of tracking on achievement in the different tracks or in the school as a whole.

How Do Track Effects Occur?

Prior research suggests that tracking influences student achievement through mechanisms of social-psychological and academic differentiation. Observers in Britain and the United States have argued that secondary school stratification polarizes the student body into pro- and anti-school factions (Hargreaves, 1967; Lacey, 1970; Metz, 1978; Ball, 1981; Schwartz, 1981). College-bound students conform to the school's demands, while others resist. The polarization process is stimulated by the labeling of students according to their track positions (Hargreaves, 1967; Lacey, 1971; Cottle, 1974; Metz, 1978; Ball, 1981; Schwartz, 1981; Burgess, 1983). These labels are said to stigmatize low-track students, inhibiting their school performance. Besides labels, teachers and guidance counselors communicate their differential expectations to students by encouraging those in college-bound programs more than others (Hargreaves, 1967; Heyns, 1974; Ball, 1981; Schwartz, 1981). Academic-track students also have more college-bound friends than students enrolled in other programs (Hargreaves, 1967; Lacey, 1970; Cottle, 1974; Hauser, Sewell, and Alwin, 1976; Rosenbaum, 1976; Ball, 1981; Alexander and Cook, 1982). As a result, high-track students are believed to find greater meaning in schoolwork, to be more motivated, to put forth greater effort, and to hold higher expectations for themselves. All this is said to lead to differences in achievement (among other outcomes).

No quantitative study has actually tested this view by placing student behavior, attitudes and expectations as intervening variables between track positions and achievement. In fact, the evidence is inconclusive as to whether tracking actually causes such social-psychological differentiation, or whether it merely reflects differences already in place. A great many studies have reported track differences in educational expectations, and several included controls for plans at the outset of tracking (Rehberg and Rosenthal, 1978; Alexander and Cook, 1982; Waitrowski et al., 1982; Vanfossen, Jones, and Spade, 1987; Berends, 1990). Thus it seems safe to conclude that tracking differentiates expectations beyond variation that existed before students were assigned to tracks. The evidence concerning student attitudes and behavior is more ambiguous. Waitrowski et al. (1982) found no track effects on self-esteem, attachment to school, or delinquent behavior, and Berends (1990) failed to support the polarization hypothesis for engagement in schoolwork, discipline problems, or absenteeism. However, Vanfossen, Jones, and Spade (1987) reported significant track effects on self-esteem and liking for school. One may conclude that tracking is implicated in the differentiation of student expectations, and possibly student attitudes, and that variation in attitudes and expectations may contribute to variation in achievement.

Besides the possibility of social-psychological differentiation, tracking appears to produce differences in students' academic experiences that further differentiate achievement. Students in college-preparatory programs take more academic courses, particularly in math and science (Gamoran, 1987; Vanfossen, Jones, and Spade, 1987). In many subject areas, they are exposed to more high-status knowledge (Keddie, 1971; Burgess, 1983, 1984; Oakes, 1985; Page, 1987). Teachers in high-track classes present more complex material at a faster pace (Metz, 1978; Ball, 1981; Oakes, 1985), and both survey and observational studies have reported a more positive academic climate in high-track classes (Metz, 1978; Oakes, 1985; Vanfossen, Jones, and Spade, 1987). Finally, teachers judged to be more skillful are disproportionately assigned to high-track classes (Hargreaves, 1967; Lacey, 1970; Rosenbaum, 1976; Ball, 1981; Finley, 1984). Although these between-track differences seem clearly documented, their role as mediators of the relation between tracking and achievement has not been well established (see Gamoran and Berends, 1987, and Gamoran, Berends, and Nystrand, 1990, for further discussion). Still, it would not be incautious to view instructional differentiation as an important mechanism through which track differences in achievement come about.

Summary of Findings about Track Effects

Prior research gives rise to four tentative conclusions:

Finding #1: Tracking adds to inequality of achievement.

Finding #2: Tracking may have no overall effect on achievement, if high-track gains are offset by low-track losses as appeared in a national study of ability grouping in Britain.

Finding #3: Tracking contributes to inequality in part through its association with variation in students' attitudes toward school and their expectations for future schooling. High-track students have higher expectations for successful performance, while low-track students tend to be more alienated from the school's demands.

Finding #4: Track differences in achievement also result from between-track variation in students' academic experiences. Teachers in high-track classes may be more skillful; they present more complex material at a faster pace; and college-track students take more advanced courses, especially in math and science.

Finding #1 is the most well-established, and Finding #2 is least secure. Even if the second is incorrect, the first underscores the importance of learning whether some tracking systems entail less inequality than others, and whether some are more productive than others. Findings #3 and #4 are essential from a theoretical standpoint, because they allow us to construct hypotheses about how track effects may vary. That is, given the mechanisms of social-psychological and instructional differentiation, one can predict the conditions under which the effects of tracking will be larger or smaller.

Student Achievement and the Structure of Stratification

With the results of prior research in mind, we are prepared to discuss the structural dimensions of high school tracking systems, and their impact on achievement inequality and productivity.

Selectivity

Sorensen (1970) described selectivity as the amount of homogeneity that educators intend to create by dividing students into groups according to characteristics relevant for learning. A selective system, in his view, is one that organizes students for instruction in classes that are more homogeneous than the student body as a whole. One can also conceive of selectivity as the size of the gap between status groups--a more selective system would be one in which the top group is much higher on the selection-criterion (e.g., ability) than other groups (Gamoran, 1984). Thus, selectivity involves both the variance (homogeneity) and the mean (level) of the high-status group relative to other groups and the total population.

By definition, highly selective track systems are elitist--they place high-achieving students together in homogeneous classes. Tracking is likely to be highly visible in more selective systems, with particularly high academic status awarded to the "cream of the crop." By emphasizing the value of the top track at the expense of other positions, selectivity is likely to add to between-track variation in students' educational attitudes and expectations. Combining this reasoning with Finding #3, one would expect high selectivity to accentuate between-track differences in achievement.

Moreover, selective tracking systems are likely to be characterized by greater between-track variability in students' instructional experiences. Because teachers adjust instruction to student aptitudes, tracks that differ more in initial levels of student performance are likely to vary more in their instructional regimes (Dahloff, 1971; Lundgren, 1972; Barr and Dreeben, 1983). Larger instructional differences would produce, if Finding #4 is correct, larger net achievement differences.

Hypothesis #1a: The greater the selectivity of a tracking system, the larger the achievement differences between tracks, controlling for relevant prior characteristics of students.

At the same time, greater selectivity may lead to higher achievement overall. When the tracks are more homogeneous, teachers can more effectively tailor the curriculum to students' needs (Sorensen, 1970). Hence, if there is an instructional advantage to homogeneous grouping, the advantage may be greater when the groups are more homogeneous (Slavin, 1987).

Hypothesis #1b: The greater the selectivity of a tracking system, the higher the overall achievement in the school, controlling for the composition of the student body.

Hypothesis #1b describes only the school average, and does not differentiate between effects on productivity in the different tracks. However, the combination of #1a and #1b implies that selectivity adds to inequality by raising achievement in the higher track faster than it raises achievement in the lower track.[1]

Electivity

Electivity refers to the extent to which students choose or are assigned to their track positions (Sorensen, 1970). Several researchers have reported that even when students are formally given their choice of track, in actuality they are led to make a selection designated by school authorities (Cicourel and Kitsuse, 1963; Ball, 1981). Students and their parents are urged by teachers, principals, and guidance counselors to make the "right" choice according to their capacities.

Nonetheless, a relatively large number of American high school students believe they chose their own track positions (Jencks et al., 1972; Jones, Vanfossen, and Spade, 1986).[2] These perceptions may actually be more significant for track effects on achievement than the objective circumstances in which assignment takes place. Students who believe they selected their programs are more likely to be motivated to perform, regardless of which track they are in. Thus, greater electivity may result in greater effort in both tracks. On the basis of Finding #3, one may expect less social-psychological differentiation in an elective tracking system, and consequently smaller between-track achievement differences. Because the lower degree of differentiation occurs through more positive attitudes in both tracks, one may also predict higher overall achievement in a more elective system.

Hypothesis #2a: The higher the degree of electivity in a tracking system, the smaller the achievement differences between tracks, controlling for relevant prior characteristics of students.

Hypothesis #2b: The higher the degree of electivity in a tracking system, the higher the average achievement in the school, controlling for the composition of the student body.

In contrast to this reasoning, Sorensen (1970) suggested that electivity leads to larger track effects on achievement. He argued that greater electivity would result in more within-track homogeneity of educational aspirations, which would produce larger track differences in achievement.

It is well known that tracking affects educational aspirations (e.g., Heyns, 1974; Hauser, Sewell, and Alwin, 1976; Rosenbaum, 1980; Alexander and Cook, 1982; Vanfossen, Jones, and Spade, 1987). But it seems to me that greater homogeneity of aspirations within tracks would result in smaller rather than larger achievement differences. Students who choose the noncollege track would be more likely to benefit from instruction and less likely to rebel than students who are assigned against their wishes.[3] Greater effort and less resistance among noncollege-track students would lead to smaller track differences, contrary to Sorensen's prediction.

My argument for the positive effects of electivity on productivity may also be challenged. The contrary position would hold that given a choice, students would elect to avoid the more challenging academic track, and consequently achieve less (Powell, Farrar, and Cohen, 1985; Kilgore, 1990). Once we control for the size of the academic track, however, electivity would be expected to exhibit a positive effect on achievement as described in Hypothesis #2b.

Inclusiveness

A more inclusive stratification system is one that tends to leave open students' options for future schooling (Sorensen, 1970; see also Rosenbaum, 1976). A tracking system is more inclusive when it assigns more students to a college-preparatory curriculum (see Kilgore, 1990, for an alternate view of inclusiveness). The larger the size of a college track, the more salient it would seem to be--for those who are left out. It seems likely that the stigma of being excluded would be greater when a larger proportion of students are included. For example, membership in the noncollege program might be more stigmatizing when it consists of the bottom 10% of the school's academic hierarchy than when it contains the bottom 40%. So although an inclusive system is less elitist, it is still highly visible and may be stigmatizing for those left out of the preferred group.

However, a system characterized by very low inclusiveness would also probably raise the salience of the college track. Like high selectivity, low inclusiveness reflects an elitist system, which can be expected to raise the degree of social-psychological and instructional differentiation corresponding to tracking. Hence, one would expect larger track differences in achievement when inclusiveness is very low as well as when it is high. One would expect smaller differences when students are more evenly distributed across tracks.

Hypothesis #3a: Controlling for relevant prior characteristics of students, track differences in achievement are larger when the system is highly inclusive or minimally inclusive, and smaller when inclusiveness is moderate.

The impact of inclusiveness on productivity may also be nonlinear. One may expect higher average achievement in schools with larger college tracks, net of composition and individual students' track positions, because a large college track reflects greater academic emphasis in the school, which may produce higher achievement for all students regardless of track (Powell, Farrar, and Cohen, 1985; Lee and Bryk, 1988). This advantage, however, probably occurs at a declining rate, because as inclusiveness becomes very high, those who are left out may become increasingly stigmatized (see hypothesis #3a), exerting a downward pull on mean achievement. Thus, as the size of the academic track increases, the benefits of inclusiveness may decline.

Hypothesis #3b: Controlling for compositional differences, inclusiveness in a tracking system contributes to average achievement, but at a declining rate.

This nonlinearity may account for the weak linear effects of the size of the academic track on student achievement observed in earlier work (Gamoran, 1987).

Scope

Sorensen (1970) viewed scope as the extent to which students are located in the same track across subjects. This conception was elaborated by Rosenbaum (1976), who included mobility--the extent of changes in students' track positions--as a characteristic of a track system.[4] Oakes (1985) further specified aspects of the scope of tracking by distinguishing among "extent" (the proportion of the total of classes that are tracked); "pervasiveness" (the number of subject areas that are tracked); and "flexibility" (whether assignments are made subject by subject or across all subjects).

Track systems with wider scope are likely to be more salient to students. Status distinctions may be more meaningful if they apply to more of a student' school day, and if they are consistent across subjects for a given student. In addition, broader and more permanent track systems are more likely to lead to differential friendship networks (Sorensen, 1970). The socialization effects of tracking would thus be compounded in a system of wider scope, and Finding #3 indicates that achievement differences would be larger.

Wider scope also means more between-track variation in students' academic experiences. Students grouped for more subjects and for a longer period of time would be exposed to a greater amount of differentiated instruction. Thus, Finding #4 would also suggest that wider scope leads to larger net track effects.

Hypothesis #4a: The wider the scope of a tracking system, the larger the achievement differences between tracks, controlling for relevant prior characteristics of students.

At the same time, a tracking system that is inflexible over time and across subjects may result in lower achievement overall. Failure to adjust assignments for developmental, motivational, or other changes in students' capacities for learning, and failure to recognize differences in a given student's aptitudes for different subjects, impede the matching of instruction to students' needs (Slavin 1987). Hence, the differentiating effect of wide scope is likely to occur in the context of lower overall achievement.

Hypothesis #4b: The wider the scope of the tracking system, the lower the average achievement in the school, controlling for student body composition.

Tracking in Public and Catholic Schools

Research on public and private schools gives reason to believe that Catholic schools use tracking differently than public schools. First, Catholic schools make greater academic demands of students in noncollege tracks, requiring more academic courses and more rigorous classwork, compared to public-school noncollege tracks (Hoffer, Greeley, and Coleman, 1985; Lee and Bryk, 1988). Hence, the degree of instructional differentiation may be lower in Catholic schools. Second, an observational study of three Catholic high schools reported that students and teachers held positive views about assignment to lower tracks and were optimistic about the possibility of remediation there (Valli, in press). This contrasted with the negative attitudes typically reported in public schools (e.g., Oakes, 1985). The contrast suggests that tracking's social-psychological differentiation may be less extensive in Catholic schools. Combined with Findings #3 and #4, the evidence suggests that net achievement gaps between tracks are likely to be smaller in Catholic than in public schools.

Hypothesis #5a: Achievement differences between tracks are smaller in Catholic schools than in public schools, controlling for relevant prior characteristics of students.

Although the finding is controversial, several studies have indicated that average achievement is higher in Catholic than in public schools (Hoffer, Coleman, and Greeley, 1985; Lee and Bryk, 1988, 1989; for critiques, see Alexander and Pallas, 1985; Willms, 1985; and Jencks, 1985). Part of the Catholic-school advantage may be tied to the way tracking is used; in particular the larger size of the academic track, the emphasis on academic work in all tracks, and the lesser stigmatization of low-track students may contribute to higher achievement in Catholic schools (Hoffer, Coleman, and Greeley, 1985; Lee and Bryk, 1988; Valli, in press). Hence, one may predict higher overall achievement in Catholic compared with public schools.

Hypothesis #5b: Catholic schools produce higher overall achievement, net of compositional differences, compared to public schools. Differences in the structure of tracking account for part of the Catholic-school advantage.

Methods

These hypotheses describe effects at two levels of analyses: (1) student-level effects on achievement within schools, and (2) school-level effects, including both conditions that affect between-school variation in the impact of tracking, and effects on variation in school mean achievement, net of compositional differences. To address the two levels of analysis, this paper uses a method called hierarchical linear modeling (HLM) (Bryk and Raudenbush, 1990), also known as multi-level contextual analysis (Mason, Wong, and Entwistle, 1983). HLM relies on equations that correspond to the two levels of analysis. At the student level, achievement is predicted within each school:

$$(achievement)_{ij} = \beta_{0j} + \beta_{1j}(track)_{ij} + \beta_{2j}(background)_{ij} + e_{ij} \quad (1)$$

In this study, both the intercept β_{0j} and the track effect β_{1j} will be allowed to vary from school to school, while the effects of background variables β_{2j} will be constrained to be equal across schools.[5] The β coefficients that vary across schools serve as dependent variables in the school-level equations:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(sector)_j + \gamma_{02}(structure)_j + \upsilon_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(sector)_j + \gamma_{12}(structure)_j + \upsilon_{1j} \quad (3)$$

When the within-school predictors are centered around their grand means, β_{0j} represents school mean achievement adjusted for compositional difference; i.e. net school productivity.[6] Inequality across tracks is reflected in β_{1j}, which is the net achievement gap between tracks in each school. In HLM, equations 1, 2, and 3 are estimated simultaneously, using the EM algorithm to obtain maximum likelihood estimates of the variance components, which are then used to generate the β and γ coefficients (for the EM algorithm, see Dempster, Laird, and Rubin, 1977; for a more detailed account of how the estimates are obtained, see Bryk and Raudenbush, 1990).

The HLM approach is superior to more traditional techniques for measuring school effects and track effects. For example, a common strategy would be to estimate the entire model at the student level, assigning the values for school-level variables to students within schools. This approach would use ordinary least squares (OLS) regression to obtain the track effects, adding interaction terms to assess the impact of sectoral and structural variation on the effects of tracking:

$$(achievement)_i = \beta_0 + \beta_1(track)_i + \beta_2(background)_i + \beta_3(sector)_i + \beta_4(structure)_i$$

$$+ \beta_5(track \times sector)_i + \beta_6(track \times structure)_i + e_i \quad (4)$$

There are several advantages to estimating Equations 1, 2, and 3 with HLM instead of relying on the OLS approach reflected in Equation 4. First, it is evident that Equation 4 contains only one error term, in contrast to Equations 1, 2, and 3, which partition error variance into its within-school and between-school components. OLS confounds the two sources of error, a problem that is particularly serious when the level-1 observations are clustered within level-2, such as when students are surveyed within schools. This violates the assumption of independent errors in the individual-level model (Equation 4), leading to underestimated standard errors (see further Goldberger and Cain, 1982). In contrast, by estimating distinct school- and student-level errors, HLM adjusts for the correlation of errors within schools (for further details see Bryk and Raudenbush, 1990).

Second, the HLM estimators of the β coefficients are more precise than those generated by OLS. In estimating the equations for both levels simultaneously, HLM uses information from all the schools to weight the β's according to their precision. Estimates that are more reliable, such as those from larger samples, are given greater weight, and the less reliable β's are "shrunk" toward their predicted values (on the "shrinkage" estimator, see Raudenbush, 1988). The β's estimated by HLM thus have smaller mean square error than those from OLS (Mason, Wong, and Entwistle, 1983; Raudenbush, 1988).

A third advantage to using HLM, particularly for this study, is that it allows us to estimate the total school-level parameter variance in β_{0j} and β_{1j}, both before and after the multi-level interactions are considered. This procedure will provide a sense of the degree to which the school structure and sector variables account for net between-school variation in achievement and track effects. It will also yield a convenient test for the heterogeneity of achievement means and distributions across schools.

Analyses for this paper were conducted using the HLM computer program (Bryk, et al., 1988). Kreft, Kim, DeLeeuw (1990) provide a comparison of HLM with other programs for multi-level analysis.

Data

For information on tracking, achievement, and school characteristics, the best available data set is High School and Beyond (HSB), a nationally representative sample of high school students collected beginning in 1980 (Jones et al, 1983). For the present analyses I used data collected from 964 public and Catholic schools in the 1980 (base year) and 1982 (first follow-up) surveys. Data were gathered from a random sample of up to 36 students in each school, for a total of 28,804 students in the public and Catholic schools. I included only schools with more than ten students responding, resulting in a loss of eleven schools. I also eliminated four schools in which all students and thirty schools in which no students reported belonging to the college-preparatory track. By drawing on information from other groups, HLM is actually capable of estimating parameters for a variable that has no within-group variance (e.g., when all students are in the same track). However, the study concerns the impact of differences in the structure of tracking, not tracking as compared to its absence, and moreover some of the structural variables were undefined when all students reported the same track, resulting in school-level missing data. Additional school-level missing data reduced the sample to 885 public and Catholic schools, or 92% of the originals. All students with data on track positions and 1982 achievement were included in the student-level sample, for a total of 20,800, or 77.3% of the students surveyed in the 885 schools. Missing values on student-level independent variables were imputed with regressions based on the variables with data present.

Tracking and Achievement Data

Students' track positions are indicated by their self-reported membership in an academic or nonacademic program in their sophomore year (1980). Reliance on sophomore-year track reports eliminates the problem of whether senior-year reports, which are also available in the data, are a response to achievement rather than a cause. It is important to bear in mind that the analysis reflects the effects of track membership at a particular point in time, rather than track experiences throughout the high school career (see Gamoran, 1987, and Gamoran and Mare, 1989, for further discussion of these issues). Effects are likely to be weaker when students shift positions frequently. Indeed, Hypothesis #4a predicts that track effects will be smaller in systems than contain more mobility between tracks.

Although they do not always correspond to school records,[7] students' perceptions of their track positions are relevant for track effects on achievement. Students' views of their own positions are likely to influence their selection of courses, an important source of track differences in math and science achievement (Gamoran, 1987). In addition, student perceptions are as least as closely tied as school records to the social-psychological mechanisms that produce track effects (Gamoran, 1987; Gamoran and Berends, 1987).

Senior-year (1982) achievement on multiple-choice tests of mathematics, reading, and vocabulary serve as three separate individual-level outcomes for this study. Heyns and Hilton (1982) reported reliabilities of .85 (part I) and .54 (part II) for the 38-item, two-part math test; of .78 for the 20-item reading test; and of .81 for the 21-item vocabulary test. The analyses begin by examining the extent to which schools vary in net mean achievement and in the net effect of membership in the academic program on achievement, and moves on to explore the sectoral and structural sources of between-school variation in these parameters.

Other Student-Level Variables

The individual-level equation (Equation 1) describes the predictors of achievement within each school. In addition to track position, it includes sophomore-year performance on all three tests; sex; minority status (black or Hispanic); and SES (a linear additive composite of mother's and father's education, father's occupation, family income, and home artifacts). These variables were included as controls to purge the estimated track differences from differences in the types of students assigned to varied tracks. Each of these predictors is known to be associated with tracking and with achievement (Gamoran, 1987; Gamoran and Mare, 1989). Table 1 provides sample means and standard deviations.

School-Level Variables

Catholic-school (as opposed to public-school) membership is indicated by a dummy variable. Other school-level variables describe the structural dimensions of tracking.

Selectivity. I constructed two indicators of selectivity. One is the initial achievement gap between tracks, computed as the difference, in each school, between the average sophomore math achievement of college-track students and that of noncollege-track students. The second indicates the proportion of variance in 1980 math achievement that lies between tracks, i.e. track homogeneity. Thus, the two indicators reflect selectivity in terms of means and distributions of achievement, corresponding to the conceptualization presented earlier. Larger achievement gaps between tracks, and more homogeneous tracks, indicate more selective tracking systems. On average, the initial gap between tracks was about 5 points, but less than 17% of the variance in achievement occurred between tracks (see Table 1).

Electivity. Electivity was computed as the percent of students in the school who said they chose their curricular program. This measure relies on students' perceptions of electivity, but as I noted earlier, students' perceptions of whether they chose their own track are probably more relevant for the tracking-achievement relation than the objective circumstances through which assignment occurs. In these data, about two-thirds of the students believed they chose their own track (see Table 1, and Jones, Vanfossen, and Spade, 1986).

Inclusiveness. Inclusiveness is indicated by the proportion of students in the academic track. To allow for the anticipated non-linear effects of inclusiveness, I also prepared a quadratic term for this variable. Inequality is expected to be greatest when inclusiveness is very high or very low; this would be indicated by a negative linear effect and a positive quadratic term. Productivity is supposed to rise with increasing inclusiveness, but at a declining rate; this pattern would be indicated by a positive linear effect and a negative quadratic term.

Scope. I created three indicators of scope. The first is a measure of agreement of students' track positions in their sophomore and senior years. This variable is a kappa statistic (Cohen, 1960), and it indicates the extent to which students tend to remain in the same track over time.[8] I used the kappa statistic, rather than simple proportions of students moving in and out of the college track, because it is independent of differences in the marginal distributions of students across tracks. A value of 1 for kappa indicates no mobility between the sophomore and senior years, whereas 0 indicates students were as likely to move as to stay. (Negative values for kappa are also possible, but they are unlikely in this situation because they would indicate a tendency for students to shift tracks more often than remaining). The kappa statistic for track immobility was computed separately for each school. Table 1 shows an average of .463, reflecting a fair amount of mobility as reported in previous work (Gamoran, 1987).

The other indicators are also kappa statistics: one for the extent to which students who reported taking honors math classes also tend to take honors English, and the last for whether students in remedial math also take remedial English. Wider-scoped tracking systems are thus indicated by higher values for track immobility, honors rigidity, and remedial rigidity.

Results

I estimated three HLM models for each dependent variable (math, reading, and vocabulary achievement). First is a baseline model, which produces estimates for the within-school equation, and for the variance components of the parameters that differ among schools (i.e., β_{0j} and β_{1j} from Equations 1-3). Recall that for this study, the effects of prior achievement and background are constrained to be equal across schools, whereas track effects and adjusted mean achievement vary from school to school. To facilitate computation and interpretation, the prior achievement and background variables are centered around their grand means, and the track variable is centered around school means (see note 6).

The second HLM model incorporates sector as a predictor of between-school differences in track effects and in school mean achievement (adjusted for compositional differences). Structural dimensions of tracking are added in the final model, which displays the conditions of track organization under which schools have higher or lower achievement levels and larger or smaller achievement gaps between tracks. Tables 2, 3, and 4 display the results.

Baseline Models

The first columns of Tables 2, 3, and 4 present the estimates for the within-school equation. The pattern of results is familiar. Except for the absence of a sex effect on vocabulary achievement (Table 3), each of the six within-school predictors exerts a significant impact on each test. On average, membership in the academic track has positive effects on achievement in each subject, reflecting increasing between-track inequality as reported in previous research.

The baseline model is thus consistent with Finding #1 from the literature. It does not speak to Findings #2, #3, or #4, nor is it our purpose to reassess them. However, the baseline model helps judge whether the five pairs of hypotheses are worth pursuing. As shown in the top panel of Table 5, the baseline model yields estimates of residual parameter variance for the two coefficients that vary between schools: the academic track effect, and school mean achievement, adjusted for the composition of the student body. The chi-square tests indicate significant variation for both coefficients in all three subjects. Given the model, both track effects and adjusted school achievement vary significantly across schools. The next step is to learn whether we can account for this variation.

Sector Effects

The second HLM model addresses the question of whether track effects and overall achievement vary between Catholic and public schools. All three subjects yield negative coefficients for sector differences in the impact of tracking, but the effect is statistically significant only in math (see Table 2, second column). In that subject, public-school college and noncollege students differ by 1.588 points, net of background and prior achievement, but the track achievement gap is only (1.588 - .649) = .939 points in Catholic schools, or about 40% smaller. The reduction in the vocabulary achievement track gap is equally substantial, but not statistically reliable (Table 4). In contrast, the impact of Catholic schools on the track difference in reading achievement is negligible (Table 3).

At the same time, all three subjects show the familiar Catholic-school advantage in adjusted mean achievement. Thus especially in math, Catholic schools produce less inequality between tracks in the context of higher achievement overall, supporting Hypotheses #5a and #5b. These differences may result in part from differences in the ways tracking is used in the two sectors.

Effects of the Organization of Tracking

The final columns of Tables 2, 3, and 4 display the impact of the structural dimensions of tracking on the track achievement gaps and on adjusted school achievement. Again the hypotheses find their greatest support in math. As predicted by Hypothesis #3a, the track gap is wider when inclusiveness is high or low, and smaller when inclusiveness is moderate. This finding is reflected in the negative linear coefficient (-3.318) and positive quadratic coefficient (4.971) for the percent of students in the academic track. Inclusiveness also affects school mean math achievement in the expected way: the positive linear effect and negative quadratic coefficient are consistent with a positive impact at a declining rate as in Hypothesis #3b. For inequality, the effects appear only in math, but for productivity the effects also occur in vocabulary, and the reading equation yields coefficients that are in the predicted direction but non-significant.

One other structural dimension has a consistent impact on inequality: track immobility, an indicator of scope. Schools in which students tend to remain in the same track over time exhibit larger gaps between tracks; the finding holds for all three subjects. Other things being equal, the achievement gap in a very rigid school (one standard deviation above the mean, or a kappa statistic of .704) would be wider than the gap in a very flexible school (one standard deviation below the mean, or kappa = .222) by about one and one-third points on the math test, over four-tenths of a point in reading, and one third of a point in vocabulary achievement. In math, track immobility also produces lower achievement overall, and rigidity in the honors program brings lower achievement in vocabulary. Thus, the data tend to support Hypotheses #4a and #4b.

Electivity has a positive effect on mean achievement in vocabulary, consistent with Hypothesis #2b, but no other significant effects. The lack of support for Hypothesis #2a may indicate that the process I described and Sorensen's (1970) contrary position both occur, in effect canceling out one another. Greater electivity may subdue potential rebellion as I argued, but it may also promote between-track differences in aspirations.

The weakest support for my conceptual framework comes from selectivity. Hypotheses #1a and #1b predicted positive effects on both inequality and productivity. The results indicated no effects on inequality. For productivity, there is a positive effect of the track achievement gap in math, but track homogeneity yields unexpected negative coefficients in all three subjects. It seemed possible that the unanticipated negative coefficients resulted from collinearity among the two indicators of selectivity, which correlate at .71. However, excluding either one from the analysis does not change the sign of the other (results not shown). In particular, track homogeneity still has a negative sign, though it is not always significant, even when track achievement gap is omitted from the analysis. Collinearity is not a satisfactory explanation for the results, but I lack a better one.

For both the dispersion and level of math achievement, the sector effects decline from the second to the third column (see Table 2). This pattern is consistent with the argument that part of the lower inequality and higher productivity of Catholic schools stems from differences in the structure of tracking. The Catholic-school advantage in mean achievement is less than half its original size in Table 2's final model. The same pattern for mean achievement occurs in reading and vocabulary, but the change is much smaller.

Given the greater support for the hypotheses in math achievement, it is not surprising to find that the math analysis was more successful in explaining between-school variation in the academic track effect. The bottom panel in Table 5 shows the extent to which residual parameter variance declines after including the school-level predictors in the model. Just over 40% of between-school variance in track effects is explained in math, compared to around 10% in the other two subjects. For adjusted school mean achievement, sectoral and structural variation account for about 23% of the variance in math, 8% in vocabulary, and 35% in reading. Significant variation in the parameters still remains.

Conclusions

In this study, the question of how tracking affects achievement elicits a more complex answer than it has in the past. Though not every hypothesis was supported, in general the analyses indicate that the effects of tracking depend on the structure of tracking. Schools with more rigid tracking systems tend to produce both more inequality and less overall achievement. In math, wider inequality also results from very high or very low inclusiveness, and in math and vocabulary, inclusiveness adds to achievement, though the rate of its contribution declines as its level rises. Electivity also raises achievement, but only in vocabulary. Finally, Catholic schools not only have higher achievement, net of measured background variables, but lower achievement gaps between tracks (most clearly in math), supporting previous speculation (Gamoran and Berends, 1987).

These findings underscore the importance of assessing contextual variation in microsocial processes. Although this study's findings are consistent with prior research for the general case--on average, belonging to the academic track is beneficial for achievement--the advantage does not appear in all schools. It is smaller in schools with more flexible and (in math) moderately inclusive tracking systems. At the same time, the study does not advocate a haphazard search for all possible sorts of contextual variation. Instead, one needs an *a priori* conceptual framework that suggests what dimensions of the context need to be examined. In this study, the framework for understanding the aggregate-level differences was built on knowledge of how the micro-level processes occur.

The results also draw attention to the value of examining tracking's effects on productivity as well as its implications for inequality. The finding of less between-track math inequality in Catholic schools, for example, would not by itself indicate whether Catholic schools were using tracking more or less successfully. The results could have occurred through lower scores in the higher track. By studying productivity as well, we learn that lower inequality occurred along with higher overall achievement, suggesting that the narrower gap in Catholic schools occurs because low-track students are brought up, not because high-track students are held down. Conversely, the results for scope suggest that inflexible track systems have more inequality along with lower average achievement; presumably conditions are especially poor in the lower tracks of such schools. The implications of the results for inclusiveness are even more complex: in math, inequality is lower when inclusiveness is moderate, but productivity is higher when inclusiveness is high. Consequently, an educator would have to choose between maximizing overall achievement in the school--usually a significant goal--and minimizing inequality.

Quantitative analysis always makes the world appear simpler than it really is. Ordinaily one strives for parsimony, so it is worth questioning whether the benefits of added complexity are worth the difficulties of conceptualization, estimation, and interpretation. In this case the enhanced theoretical understanding and the potential policy benefits seem to justify the effort.

Though the study made progress in accounting for between-school differences in the effects of tracking, significant variation remains, and it is worth considering other school-level influences. Future investigations might examine the culture or ethos of tracking in different schools. Tracking does not carry the same meaning in all schools: in some schools it is a clear symbol of students' future directions, but elsewhere its significance is more vague (Gamoran and Berends, 1987). Presumably, tracking's effects would be heightened in schools in which its power to confer status is greater.

NOTES

1 Sorensen (1970) also noted that schools differ in the criteria used to assign students to programs. A key issue is the extent to which placement relies on cognitive characteristics (e.g., ability or achievement). This concern appears to be subsumable under selectivity, because when the indicator of selectivity relies on achievement data, more selective systems by definition involve tighter links between cognitive characteristics and track positions.

2 Data reported by Jones, Vanfossen, and Spade (1986) from a random subsample of the nationally representative High School and Beyond survey indicate that about two-thirds of high-school sophomores said they chose their curricular program.

3 This argument is consistent with Stinchcombe's (1964) that high school students rebel when their school programs do not match their anticipated futures.

4 In a footnote, Sorensen distinguished scope from "the rigidity of differentiation...the extent to which students may transfer to another group than the one originally assigned to" (1970, p.363, note 2). Although Sorensen believed this would involve few students, recent data suggest transfers occur widely, at least as indicated by self-report data (Gamoran, 1987). It seems reasonable to consider the permanence of assignments, or track immobility, as part of tracking's scope.

5 In Equation 1, β_{2j} actually represents a number of separate background variables; I have written the equation as if there were only one for the sake of simplicity. In principle, the effects of one or more background variables could also be allowed to vary between schools. However, freeing more slopes multiplies the number of variances and covariances that need to be estimated, dramatically increasing the complexity of the model and the difficulty of estimation. For this reason, HLM users are advised to start small, freeing parameters when there is theoretical interest in their variability (Bryk and Raudenbush, 1990).

6 To adjust the within-school intercepts for variation in effects permitted to vary across schools, it is necessary to center the within-school variable around the mean for each school, and to include the school average for the variable in the equation for the intercept (Equation 2) (Bryk and Raudenbush, 1990). In this study, background variables have constant effects across schools, so they are centered around their grand means in the analyses. Because the effects of tracking are allowed to vary across schools, this variable is centered around school means and, in the later stages of the analysis, school mean track (i.e., the proportion of students in the academic track) is included as a predictor of net school achievement.

7 Data from the National Longitudinal Survey indicated that 80% of the students reported the same track position as that reported for them by school officials (Vanfossen, Jones, and Spade, 1987; see also Rosenbaum, 1980).

8 The formula for kappa is (Po - Pc) / (1 - Pc) where Po is the proportion observed and Pc is the proportion expected by chance (Agresti, 1990). For example, track immobility in a school is computed as:

$$\frac{[(PoAc + PoNAc) - (PAc80 \times PAc82 + PNAc80 \times PNAc82)]}{[1 - (PAc80 \times PAc82 + PNAc80 \times PNAc82)]}$$

where PoAc is the proportion in the academic track in both years; PoNAc is the proportion in the non-academic track in both years; PAc80 and PAc82 are the proportions in the academic track in 1980 and 1982; and PNAc80 and PNAc82 are the proportions in the non-academic track in 1980 and 1982. (Multiplying and summing the marginals as indicated yields the cell proportions expected by chance.)

REFERENCES

Alexander, K. L., and Cook, M. A. (1982). Curricula and coursework: A surprise ending to a familiar story. American Sociological Review, 47, 626-640.

Ball, S. J. (1981). Beachside Comprehensive: A case-study of secondary schooling. Cambridge: Cambridge University Press.

Barr, R., and R. Dreeben. (1983). How schools work. Chicago: University of Chicago Press.

Bryk, A. S., & Driscoll, M. E. (1988). The high school as community: Contextual influences, and consequences for students and teachers. Madison, WI: National Center on Effective Secondary Schools.

Bryk, A. S., and Raudenbush, S. W. (1990). Hierarchical linear models for social and behavioral research: Applications and data analysis methods. Chicago: Department of Education, University of Chicago.

Bryk, A. S., Raudenbush, S. W., Seltzer, M., and Congdon, R. T. (1988). An introduction to HLM: Computer program and users' guide. Version 2.0. Chicago: Department of Education, University of Chicago.

Bryk, A. S., and Thum, Y. M. (1989). The effects of high school organization on dropping out: An eploratory investigation. American Educational Research Journal, 26, 353-384.

Burgess, R. G. (1983). Experiencing comprehensive education: A study of Bishop McGregor School. London: Methuen and Co.

Burgess, R. G. (1984). It's not a proper subject: It's just Newsom. In I. Goodson and S. Ball (eds.), Defining the curriculum. London: The Falmer Press.

Cicourel, A. V., and Kitsuse, J. I. (1963). The educational decision-makers. Indianapolis: The Bobbs Merrill Co.

Cottle, T. J. (1974). What tracking did to Ollie Taylor. Social Policy, 5, 21-24.

Dahloff, U. (1971). Ability grouping, content validity, and curriculum process. New York: Teachers College Press.

Dempster, A. P., Laird, N. P., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). Journal of the Royal Statistical Society (Series B), 39, 1-8.

Finley, M. K. (1984). Teachers and tracking in a comprehensive high school. Sociology of Education, 57, 233-243.

Gamoran, A. (1987). The stratification of high school learning opportunities. Sociology of Education, 60, 135-155.

Gamoran, A. (1990). Access to excellence: Assignment to honors English classes in the transition from middle to high school. Paper presented at the annual meeting of the American Sociological Association, Washington, DC.

Gamoran, A., and Berends, M. (1987). The effects of stratification in secondary schools: synthesis of survey and ethnographic research. Review of Education Research, 57, 415-435.

Gamoran, A., Berends, M., and Nystrand, M. (1990). Classroom instruction and the effects of ability grouping: A structural model. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Gamoran, A., and Mare, R. D. (1989). Secondary school tracking and educational inequality: reinforcement, compensation, or neutrality? American Journal of Sociology, 94, 1146-1183.

Garet, M. S., and DeLany, B. (1988). Students, courses, and stratification. Sociology of Education, 61, 61-77.

Goldberger, A. S., and Cain, G. G. (1982). The causal analysis of cognitive outcomes in the Coleman, Hoffer, and Kilgore Report. Sociology of Education, 55, 103-122.

Hargreaves, D. H. (1967). Social relations in a secondary school. London: Tinling.

Hauser, R. M., Sewell, W. H., and Alwin, D. F. (1976). High school effects on achievement. In W. H. Sewell, R. M. Hauser, and D. L. Featherman (eds.), Schooling and achievement in American Society. New York: Academic Press.

Heyns, B. (1974). Social selection and stratification within schools. American Journal of Sociology, 79, 1434-1451.

Heyns, B., and Hilton, T. L. (1982). The cognitive tests for High School and Beyond: An assessment. Sociology of Education, 55, 89-102.

Hoffer, T., Greeley, A. M., and Coleman, J. S. (1985). Achievement growth in public and catholic schools. Sociology of Education, 58, 74-97.

Jencks, C. L., and Brown, M. (1975). The effects of high schools on their students. Harvard Educational Review, 45, 273-324.

Jencks, C. L., M. Smith, H. Acland, M. J. Bane, D. K. Cohen, H. Gintis, B. Heyns, and S. Michaelson. (1972.) Inequality: A reassessment of the effects of family and schooling in America. New York: Basic Books.

Jones, C., M. Clark, G. Mooney, H. McWilliams, I. Crawford, B. Stephenson, and R. Tourangeaou. (1983). High School and Beyond 1980 Sophomore Cohort First Follow-up (1982) Data File user's manual. Washington, DC: National Center for Education Statistics.

Jones, J. D., Vanfossen, B. E., and Spade, J. Z. Curriculum placement: Individual and school effects using the High School and Beyond data. Paper presented at the annual meeting of the American Sociological Association, Washington, DC.

Keddie, N. (1971) Classroom knowledge. Pp. 133-160 in M.F.D. Young (ed.), Knowledge and control. London: Collier-Macmillan.

Kerckhoff, A. C. (1986). Effects of ability grouping in British secondary schools. American Sociological Review, 51, 842-858.

Kilgore, S. (1990). Tracking in organizational context. Atlanta: Department of Sociology, Emory University.

Kreft, I. G. G., Kim, K.. and DeLeuuw, J. (1990). Comparing four different statistical packages for hierarchical linear regression: GENMOD, HLM, ML2, and VARCL. Los Angeles: Departments of Psychology and Mathematics Social Statistics Program, University of California-Los Angeles.

Lacey, C. (1970). Hightown Grammar. Manchester: Manchester University Press.

Lee, V. E., and Bryk, A. S. (1988). Curriculum tracking as mediating the social distribution of high school achievement. Sociology of Education, 62, 78-94.

Lee, V. E., and Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. Sociology of Education, 62, 172-192.

Lucas, S. R. (1990). Course-based indicators of curriculum track locations. Unpublished M.S. thesis, University of Wisconsin, Madison.

Lundgren, U. P. (1972). Frame factors and the teaching process. Stockholm: Almqvist and Wiksell.

Mason, W. M., Wong, G. Y., and Entwistle, B. (1983). Contextual analysis through the multilevel linear model. Pp. 72-103 in S. Leinhardt (ed.), Sociological Methodology 1983-84. San Francisco: Jossey-Bass.

Metz, M. H. (1978). Classrooms and corridors: The crisis of authority in desegregated secondary schools. Berkeley, California: University of California Press.

Natriello, G., Pallas, A., and Alexander, K. On the right track? Curriculum and academic achievement. Sociology of Education, 62, 109-118.

Oakes, J. (1985). Keeping track: How schools structure inequality. New Haven, CT: Yale University Press.

Page, R. N. (1987). Lower-track classes at a college-preparatory school: A caricature of educational encounters. In G. Spindler and L. Spindler (eds.), Interpretive ethnography of schooling: At home and abroad. Hillsdale, NJ: Erlbaum.

Powell, A. Farrar, E., and Cohen, D. K. (1985). The shopping-mall high school. Boston: Houghton-Mifflin.

Raudenbush, S. W. (1988). Applications of hierarchical linear models in educational research: A review. Journal of Educational Statistics, 13, 85-116.

Raudenbush, S. W., and Bryk, A. S. (1986). A hierarchical model for studying school effects. Sociology of Education, 59, 1-17.

Rehberg, R. A., and Rosenthal, E. R. (1978). Class and merit in the American high school. New York: Longman.

Rosenbaum, J. E. (1976). Making inequality. New York: Wiley.

Rosenbaum, J. E. (1980). Track misperceptions and frustrated college plans: An analysis of the effects of tracks and track perceptions in the National Longitudinal Survey. Sociology of Education, 53, 74-88.

Rosenbaum, J. E. (1984). The social organization of instructional groups. Pp. 53-68 in Peterson, P. L., Wilkerson, L. C., and and Hallinan, M. T. (eds.), The social context of instruction. Orlando, FL: Academic Press.

Schwartz, F. (1981). Supporting or subverting learning: Peer group patterns in four tracked schools. Anthropology and Education Quarterly, 12, 99-121.

Shavit, Y., & Featherman, D. L. (1988). Schooling, tracking, and teenage intelligence. Sociology of Education, 61, 42-51.

Slavin, R. E. (1987). Ability grouping and achievement in elementary schools: A best-evidence synthesis. Review of Educational Research, 57, 293-336.

Sorensen, A. B. (1970). Organizational differentiation of students and educational opportunity. Sociology of Education, 43, 355-376.

Stinchcombe, A. L. (1964). Rebellion in a high school. Chicago: Quadrangle Books.

Valli, L. (In press). A curriculum of effort: Tracking students in a Catholic high school. In R. Page and L. Valli (eds.), Curriculum differentiation: Interpretive studies in U. S. secondary schools. Albany: SUNY Press.

Vanfossen, B. E., Jones, J. D., and Spade, J. Z. (1987). Curriculum tracking and status maintenance. Sociology of Education, 60, 104-122.

Waitrowski, M. D., Hansell, S., Massey, C. R., and Wilson, D. L. (1982). Curriculum tracking and delinquency. American Sociological Review, 47, 151-160.

Table 1. Means and standard deviations of variables.

	Mean	Standard Deviation
STUDENT-LEVEL VARIABLES (n=20,800)		
1980 Math Achievement	18.911	7.119
1982 Math Achievement	20.145	8.121
1980 Reading Achievement	9.168	3.759
1982 Reading Achievement	10.112	4.144
1980 Vocabulary Achievement	10.939	4.230
1982 Vocabulary Achievement	12.647	4.589
Sex (1=female)	.520	.500
Ethnicity (1=Black or Hispanic)	.300	.458
SES	-.064	.721
Academic Track	.354	.478
SCHOOL-LEVEL VARIABLES (n=885)		
Sector		
Catholic	.088	.284
Selectivity		
Achievement Gap	5.016	4.011
Track Homogeneity	.168	.158
Electivity		
Percent Chose Own Track	.666	.174
Inclusiveness		
Percent Academic Track	.339	.190
(Percent Academic Track)2	.151	.166
Scope		
Track Immobility	.463	.241
Honors Rigidity	.440	.234
Remedial Rigidity	.528	.222

Table 2. HLM analyses of between-school differences in the effects of tracking on math achievement: Gamma coefficients, with standard errors in parentheses. N = 885 schools, 20,800 students.

Dependent Variable - 1982 Math Achievement

PREDICTOR VARIABLES Within-School Equation	Baseline Model	Sector Effects	Organizational Effects
1980 Math Achievement	.706*** (.006)	.705*** (.006)	.702*** (.007)
1980 Reading Achievement	.190*** (.013)	.189*** (.013)	.187*** (.013)
1980 Vocabulary Achievement	.149*** (.012)	.145*** (.011)	.139*** (.011)
Sex (1=female)	-.610*** (.064)	-.625*** (.064)	-.644*** (.064)
Ethnicity (1=Black or Hispanic)	-.892*** (.082)	-.934*** (.082)	-.942*** (.082)
SES	.798*** (.051)	.774*** (.051)	.705*** (.051)
Academic Track	1.508*** (.082)	1.588*** (.086)	.286 (.492)
Intercept (Net school achievement)	20.119*** (.045)	20.016*** (.046)	18.988*** (.245)
Between-School Equations			
EFFECTS ON BETWEEN-TRACK INEQUALITY:			
Sector Catholic		-.649* (.251)	-.405 (.303)
Selectivity Achievement Gap			-.028 (.032)
Track Homogeneity			.226 (.757)
Electivity Percent Chose Own Track			.635 (.508)
Inclusiveness Percent Academic Track			3.318 (1.836)
(Percent Academic Track)2			4.971* (2.079)
Scope Track Immobility			2.794*** (.393)
Honors Rigidity			.260 (.331)
Remedial Rigidity			-.088 (.347)
EFFECTS ON ADJUSTED SCHOOL PRODUCTIVITY:			
Sector Catholic		1.065*** (.146)	.595** (.180)
Selectivity Achievement Gap			.052** (.016)
Track Homogeneity			-1.188** (.419)
Electivity Percent Chose Own Track			.097 (.277)
Inclusiveness Percent Academic Track			5.097*** (.835)
(Percent Academic Track)2			-3.757*** (.938)
Scope Track Immobility			-.700** (.209)
Honors Rigidity			-.224 (.182)
Remedial Rigidity			.378 (.193)

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 3. HLM analyses of between-school differences in the effects of tracking on reading achievement: Gamma coefficients, with standard errors in parentheses. N = 885 schools, 20,800 students.

Dependent Variable - 1982 Reading Achievement

PREDICTOR VARIABLES Within-School Equation	Baseline Model	Sector Effects	Organizational Effects
1980 Math Achievement	.110*** (.004)	.110*** (.004)	.107*** (.004)
1980 Reading Achievement	.423*** (.008)	.424*** (.008)	.423*** (.008)
1980 Vocabulary Achievement	.220*** (.007)	.218*** (.007)	.216*** (.007)
Sex (1=female)	.020 (.037)	.014 (.037)	.012 (.037)
Ethnicity (1=Black or Hispanic)	-.443*** (.047)	-.465*** (.047)	-.455*** (.048)
SES	.238*** (.030)	.226*** (.030)	.215*** (.030)
Academic Track	.495*** (.047)	.505*** (.049)	.046 (.288)
Intercept (Net school achievement)	10.102*** (.024)	10.062*** (.025)	9.899*** (.137)
Between-School Equations			
EFFECTS ON BETWEEN-TRACK INEQUALITY:			
Sector Catholic		-.020 (.143)	.297 (.177)
Selectivity Achievement Gap			-.018 (.018)
Track Homogeneity			.077 (.443)
Electivity Percent Chose Own Track			.366 (.297)
Inclusiveness Percent Academic Track			-1.204 (1.075)
(Percent Academic Track)2			.668 (1.217)
Scope Track Immobility			.918*** (.230)
Honors Rigidity			.123 (.194)
Remedial Rigidity			.201 (.203)
EFFECTS ON ADJUSTED SCHOOL PRODUCTIVITY:			
Sector Catholic		.403*** (.078)	.380*** (.099)
Selectivity Achievement Gap			.009 (.009)
Track Homogeneity			-.546* (.234)
Electivity Percent Chose Own Track			.250 (.154)
Inclusiveness Percent Academic Track			.451 (.463)
(Percent Academic Track)2			-.307 (.520)
Scope Track Immobility			-.012 (.116)
Honors Rigidity			-.051 (.101)
Remedial Rigidity			-.067 (.107)

* p < .05   ** p < .01   *** p < .001

Table 4. HLM analyses of between-school differences in the effects of tracking on vocabulary achievement: Gamma coefficients, with standard errors in parentheses. N = 885 schools, 20,800 students.

Dependent Variable - 1982 Vocabulary Achievement

PREDICTOR VARIABLES Within-School Equation	Baseline Model	Sector Effects	Organizational Effects
1980 Math Achievement	.066***	.066***	.065***
	(.004)	(.004)	(.004)
1980 Reading Achievement	.206***	.206***	.206***
	(.008)	(.008)	(.008)
1980 Vocabulary Achievement	.538***	.536***	.532***
	(.007)	(.007)	(.007)
Sex (1=female)	.094*	.083*	.071
	(.039)	(.039)	(.039)
Ethnicity (1=Black or Hispanic)	-.782***	-.823***	-.790***
	(.049)	(.049)	(.049)
SES	.488***	.473***	.436***
	(.031)	(.031)	(.031)
Academic Track	.520***	.551***	.288
	(.048)	(.051)	(.297)
Intercept (Net school achievement)	12.632***	12.550***	11.851***
	(.027)	(.027)	(.143)
Between-School Equations			
EFFECTS ON BETWEEN-TRACK INEQUALITY:			
Sector			
Catholic		-.227	-.246
		(.147)	(.182)
Selectivity			
Achievement Gap			-.024
			(.019)
Track Homogeneity			-.037
			(.456)
Electivity			
Percent Chose Own Track			-.486
			(.306)
Inclusiveness			
Percent Academic Track			1.347
			(1.107)
(Percent Academic Track)2			-1.420
			(1.253)
Scope			
Track Immobility			.704**
			(.236)
Honors Rigidity			.324
			(.200)
Remedial Rigidity			.008
			(.209)
EFFECTS ON ADJUSTED SCHOOL PRODUCTIVITY:			
Sector			
Catholic		.838***	.700***
		(.086)	(.104)
Selectivity			
Achievement Gap			.023*
			(.009)
Track Homogeneity			-.693**
			(.244)
Electivity			
Percent Chose Own Track			.653***
			(.161)
Inclusiveness			
Percent Academic Track			2.212***
			(.485)
(Percent Academic Track)2			-1.409*
			(.544)
Scope			
Track Immobility			-.127
			(.122)
Honors Rigidity			-.319**
			(.106)
Remedial Rigidity			-.124
			(.112)

* p < .05 ** p < .01 *** p < .001

Table 5. Chi-square tests for homogeneity of parameter variance.

BASELINE MODEL

Effect	Residual Paramenter Variance	Degrees of Freedom	Chi-square	p-value
MATH				
Track effect	.727	874	7144	<.001
Net school achievement	.888	874	14764	<.001
READING				
Track effect	.150	874	4901	<.001
Net school achievement	.195	874	8942	<.001
VOCABULARY				
Track effect	.159	874	5339	<.001
Net school achievement	.316	874	13905	<.001

SECTOR EFFECTS MODEL

Effect	Residual Paramerter Variance	Degrees of Freedom	Chi-square	p-value	Net Variance Explained
MATH					
Track effect	.697	874	7063	<.001	4.1%
Net school achievement	.790	873	14304	<.001	11.1%
READING					
Track effect	.151	874	4879	<.001	0.0%
Net school achievement	.181	873	873	<.001	7.2%
VOCABULARY					
Track effect	.155	874	5292	<.001	2.4%
Net school achievement	.251	873	13189	<.001	20.7%

ORGANIZATIONAL EFFECTS MODEL

Effect	Residual Paramenter Variance	Degrees of Freedom	Chi-square	p-value	Net Variance Explained
MATH					
Track effect	.430	874	6638	<.001	40.8%
Net school achievement	.681	865	13138	<.001	23.3%
READING					
Track effect	.135	874	4766	<.001	9.4%
Net school achievement	.180	865	8435	<.001	7.7%
VOCABULARY					
Track effect	.140	874	5220	<.001	11.9%
Net school achievement	.206	865	11813	<.001	34.9%